

**Perbandingan Analisis Faktor Penyalahgunaan Narkoba
Menggunakan Metode *Naïve Bayes*
dan *Decision Tree***



**Disusun sebagai salah satu syarat menyelesaikan Jenjang Strata I pada
Jurusan Informatika Fakultas Komunikasi dan Informatika**

Oleh:

BADRIYAH MEI JAYANTI

L 200 130 123

**PROGRAM STUDI INFORMATIKA
FAKULTAS KOMUNIKASI DAN INFORMATIKA
UNIVERSITAS MUHAMMADIYAH SURAKARTA
2017**

HALAMAN PERSETUJUAN

**PERBANDINGAN ANALISIS FAKTOR PENYALAHGUNAAN
NARKOBA MENGGUNAKAN METODE *NAIVE BAYES*
DAN *DECISION TREE***

PUBLIKASI ILMIAH

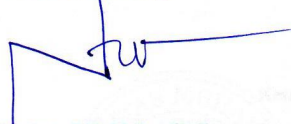
oleh:

BADRIYAH MEI JAYANTI

L 200 130 123

Telah diperiksa dan disetujui untuk diuji oleh:

Dosen Pembimbing



Nurgiyatna, S.T., M.Sc., Ph.D.

NIK.881

HALAMAN PENGESAHAN

**PERBANDINGAN ANALISIS FAKTOR PENYALAHGUNAAN
NARKOBA MENGGUNAKAN METODE *NAIVE BAYES*
DAN *DECISION TREE***

OLEH:

BADRIYAH MEI JAYANTI

L200130123

Telah dipertahankan di depan Dewan Penguji
Fakultas Komunikasi dan Informatika
Universitas Muhammadiyah Surakarta
Pada hari Senin, 17 April 2017
dan dinyatakan telah memenuhi syarat

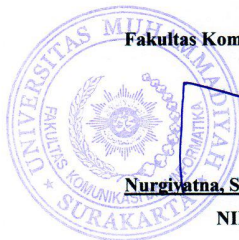
Dewan Penguji:

1. Nurgiyatna, S.T., M.Sc., Ph.D.
(Ketua Dewan Penguji)
2. Drs. Sujalwo, M.Kom
(Anggota 1 Dewan Penguji)
3. Aris Rakhmadi, S.T., M.Eng.
(Anggota 2 Dewan Penguji)

(.....)
(.....)
(.....)

Dekan

Fakultas Komunikasi dan Informatika

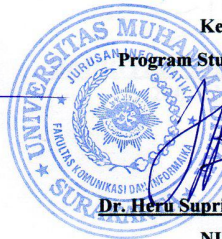


Nurgiyatna, S.T., M.Sc., Ph.D.

NIK. 881

Ketua

Program Studi Informatika



Dr. Heru Suprivono, S.T., M.Sc.

NIK. 970



PERNYATAAN

PERALIHAN KELOMPOK LULUSAN

15.03.2017-PKI/VIII/2017

Dengan ini saya menyatakan bahwa dalam naskah publikasi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan orang lain, kecuali secara tertulis diacu dalam naskah dan disebutkan dalam daftar pustaka.

Apabila kelak terbukti ada ketidakbenaran dalam pernyataan saya di atas, maka akan saya pertanggungjawabkan sepenuhnya.

Program Studi : Informatika
Status : Lulus

Surakarta, Agustus 2017

Penulis

BADRIYAH MEI JAYANTI
L 200 130 123

Surakarta, 9 Agustus 2017

Surakarta, 9 Agustus 2017





UNIVERSITAS MUHAMMADIYAH SURAKARTA
FAKULTAS KOMUNIKASI DAN INFORMATIKA
PROGRAM STUDI INFORMATIKA

Jl. A Yani Tromol Pos 1 Pabelan Kartasura Telp. (0271)717417, 719483 Fax (0271) 714448
Surakarta 57102 Indonesia. Web: <http://informatika.ums.ac.id>. Email: informatika@ums.ac.id

SURAT KETERANGAN LULUS PLAGIASI

253/A.3-IL.3/INF-FKI/VIII/2017

Assalamu'alaikum Wr. Wb

Biro Tugas Akhir Program Studi Informatika menerangkan bahwa :

Nama : BADRIYAH MEI JAYANTI
NIM : L200130123
Judul : PERBANDINGAN ANALISIS FAKTOR PENYALAHGUNAAN
NARKOBA MENGGUNAKAN METODE NAÏVE BAYES DAN
DECISION TREE

Program Studi : Informatika
Status : **Lulus**

Adalah benar-benar sudah lulus pengecekan plagiasi dari Naskah Publikasi Tugas Akhir,
dengan menggunakan aplikasi Turnitin.

Demikian surat keterangan ini dibuat agar dipergunakan sebagaimana mestinya.

Wassalamu'alaikum Wr. Wb

Surakarta, 9 Agustus 2017

Biro Tugas Akhir Informatika



Endang Wahyu Pamungkas, S.Kom., M.Kom.

Feedback Studio - Google Chrome

turnitin.com/app/carta/en_us/?o=83599059&u=1057550080&lang=en_us&s=1

PERBANDINGAN ANALISIS FAKTOR PENYALAHGUNAAN NARKOBA MENGGUNAKAN METODE NAÏVE BAYES DAN DECISION TREE

17

13 of 35

Match Overview

17%

1eprints.ums.ac.idInternet Source9%>

2Submitted to Universita...Student Paper3%>

3digilib.unimed.ac.idInternet Source1%>

4www.sersc.orgInternet Source1%>

5www.testkitshop.comInternet Source1%>

6Submitted to Universita...Student Paper1%>

7journal.ui.ac.idInternet Source1%>

8mardoto.comInternet Source<1%>

9azzaivana.blogspot.comInternet Source<1%>

Activate Windows
Go to PC settings to activate Windows.

PERBANDINGAN ANALISIS FAKTOR PENYALAHGUNAAN NARKOBA
MENGUNAKAN METODE NAÏVE BAYES
DAN DECISION TREE

Abstrak

Narkoba merupakan zat adiktif yang biasa di**3**akai dibidang kesehatan. Akan tetapi, penyalahgunaan narkotika dan obat-obatan terlarang tersebut **kian** meningkat. Maraknya penyimpanan perilaku oleh generasi muda tersebut dapat merusak masa depan dan berlangsungnya kehidupan menjadi terganggu di waktu yang akan datang. Selubungan dengan terjadinya masalah tersebut dilakukan teknik *data mining* untuk mencari faktor yang mempengaruhi lama penggunaan seseorang mengkonsumsi narkoba dan dapat diupayakan melakukan penyuluhan pada faktor yang paling berpengaruh. Metode yang dipilih yaitu *Naive Bayes* dan *Decision Tree* algoritma *Index Gini*. Hasil dari penelitian memperlihatkan prediksi *naive bayes* yaitu nilai *accuracy* 65%, *precision* 65.66%, *recall* 98.48%. Hasil *index gini* memiliki nilai *accuracy* 60%, *precision* 65.60%, *recall* 82.83%. Setelah melihat hasil tersebut disimpulkan *naive bayes* lebih akurat digunakan dalam penelitian ini karena nilai akurasi lebih besar dibanding *decision tree* algoritma *index gini*. Dari hasil pohon keputusan diperoleh variabel yang paling berpengaruh dalam penelitian ini yaitu faktor lingkungan.

Kata Kunci: Narkoba, *Data Mining*, *Naive Bayes*, *Index Gini*.

Abstract

Narcotics are addictive substances commonly used in the field of health. However, the abuse of narcotics

Page: 1 of 11

Word Count: 2039

PERBANDINGAN ANALISIS FAKTOR PENYALAHGUNAAN NARKOBA MENGUNAKAN METODE *NAÏVE BAYES* DAN *DECISION TREE*

Abstrak

Narkoba merupakan zat adiktif yang biasa dipakai dibidang kesehatan. Akan tetapi, penyalahgunaan narkotika dan obat-obatan terlarang tersebut kian meningkat. Maraknya penyimpangan perilaku oleh generasi muda tersebut dapat merusak masa depan dan berlangsungnya kehidupan menjadi terganggu di waktu yang akan datang. Sehubungan dengan terjadinya masalah tersebut dilakukan teknik *data mining* untuk mencari faktor yang mempengaruhi lama penggunaan seseorang mengkonsumsi narkoba dan dapat diupayakan melakukan penyuluhan pada faktor yang paling berpengaruh. Metode yang dipilih yaitu *Naive Bayes* dan *Decision Tree* algoritma *Index Gini*. Hasil dari penelitian memperlihatkan prediksi *naive bayes* yaitu nilai *accuracy* 65%, *precision* 65,66%, *recall* 98,48%. Hasil *index gini* memiliki nilai *accuracy* 60%, *precision* 65,60%, *recall* 82,83%. Setelah melihat hasil tersebut disimpulkan *naive bayes* lebih akurat digunakan dalam penelitian ini karena nilai akurasi lebih besar dibanding *decision tree* algoritma *index gini*. Dari hasil pohon keputusan diperoleh variabel yang paling berpengaruh dalam penelitian ini yaitu faktor lingkungan.

Kata Kunci: Narkoba, *Data Mining*, *Naive Bayes*, *Index Gini*.

Abstract

Narcotics are addictive substances commonly used in the field of health. However, the abuse of narcotics and drugs is increasing. The rise of behavioral deviation by the younger generation can ruin the future and the ongoing life becomes disrupted in the future. In connection with the occurrence of these problems is done data mining techniques to find the factors that affect the long use of someone taking drugs and can be attempted to counsel on the most influential factors. The selected method is Naive Bayes and Decision Tree Gini Index algorithm. The results of the study showed naive bayes predictions of 65% accuracy, precision 65.66%, recall 98.48%. The result of this index has accuracy value of 60%, precision 65,60%, recall 82,83%. After seeing the results concluded naive bayes more accurately used in this study because the value of accuracy is greater than the decision tree algorithm index gini. From the decision tree results obtained the most influential variables in this study that is environmental factors.

Keywords: Narcotics, *Data Mining*, *Naive Bayes*, *Gini Index*.

1. PENDAHULUAN

Narkoba dalam bidang kesehatan merupakan senyawa-senyawa psikotropika yang biasa dipakai untuk membius pasien ketika akan dioperasi atau obat-obatan untuk penyakit tertentu. Akan tetapi, peredaran mengenai narkoba saat ini luas diIndonesia. Dalam survei yang dilakukan Badan Narkotika Nasional (BNN) dengan Pusat

Penelitian Kesehatan Universitas Indonesia tahun 2008, kian meningkat penggunaan narkoba. Menurut perkiraan BNN, volume perdagangan tercatat Rp 15,4 triliun. Kenaikan tersebut terjadi pada tahun 2013 sebesar 1,89% dari jumlah penduduk. (Ricardo. 2010).

Iswanto, dkk (2015) melakukan teknik *data mining* untuk diagnosis penyakit Tuberculosis (TBC). Hasil penelitian menjelaskan teknik pengolahan data dari beberapa data rekam medik pasien yang dilakukan dengan pengambilan keputusan dalam mendiagnosis terjadinya penyakit TBC pada pasien. Teknik yang dilakukan yaitu preprocessing data dan data mining untuk diagnosis penyakit TBC. Tujuan penelitian tersebut adalah studi literatur tentang pemanfaatan teknik penambangan data untuk mendiagnosis penyakit Tuberculosis (TBC).

Hasilnya menunjukkan bahwa atribut yang paling sering digunakan yaitu: keringat malam, batuk lebih dari 3 minggu, demam, berat badan turun, usia dan nyeri dada. Support Vector Machine dan Bayesian Network memberikan akurasi tertinggi dibandingkan dengan metode lainnya. Hasil dari beberapa gejala tersebut menunjukkan klasifikasi yang akurat yaitu dengan menggunakan teknik data mining, sehingga dapat membantu dokter paru untuk pengambilan keputusan dengan cepat dan tepat.

Data mining adalah suatu metode yang digunakan untuk mengidentifikasi informasi pada sebuah *database* yang besar. Metode ini menggali data yang tersembunyi pada sebuah *database* sehingga menghasilkan pengetahuan baru dari data yang disimpan (Nasution, 2014). Dalam penelitian ini metode *naive bayes*, *decision tree* algoritma *index gini* dijadikan metode untuk mencari metode yang lebih baik digunakan dalam penelitian.

Penelitian lain dilakukan oleh Al-Bayaty dan Joshi (2016) yang menganalisis perbandingan antara algoritma *naive bayes* dan *decision tree* dalam memecahkan masalah WSD dengan menggunakan Pendekatan Empiris. Hasil dari penelitian tersebut menyatakan bahwa tidak ada metode yang nilai keakuratannya mencapai 100%, karena hal ini tergantung pada sekelompok data dan algoritma yang digunakan. Akan tetapi dalam penelitian ini, nilai *accuracy* cenderung bervariasi dan metode *naive bayes* memberikan hasil yang lebih akurat dalam waktu yang sama dibandingkan dengan metode *decision tree*.

2. METODE

2.1. Penentuan Atribut dan Variabel

Atribut dan variabel yang dipilih untuk memproses data sebagai berikut :

Tabel 1. Atribut Faktor Penyalahgunaan Narkoba

Atribut	Variabel
Jenis Kelamin	X1
Pendidikan	X2
Kepribadian	X3
Lingkungan	X4
Penggunaan Narkoba	Y

2.2. Pengumpulan Data

Data pemakai narkoba di Panti Rehabilitasi Sinai Sukoharjo dijadikan sebuah data yang dipilih berdasarkan catatan rekam medis pemakai sebanyak 331 data pemakai.

2.3. Analisis Data

Tahap ini menganalisa data yang sudah didapatkan dengan metode yang sudah ditentukan kemudian disusun dalam bentuk tabel-tabel untuk selanjutnya hasil data yang diperoleh dianalisis dan dikelompokkan.

2.4. Penggunaan *Naïve bayes*

Bayes adalah pengklasifikasian yang memprediksi nilai probabilitas suatu kelas. Rumus *naive bayes* yang dinyatakan oleh Nugroho dan Haryati (2015) terdapat pada persamaan 1.

$$P(Y | X) = \frac{P(Y | X)P(X)}{P(X)} \quad (1)$$

Keterangan:

X = Data dimana kelas belum diketahui

Y = Hipotesis kelas spesifik

P(Y|X) = Probabilitas hipotesis pada kondisi X

$P(Y)$ = Probabilitas hipotesis Y
 $P(X|Y)$ = Probabilitas X sesuai hipotesis Y
 $P(X)$ = Probabilitas X

2.5. Penggunaan *Decision Tree* Algoritma *Index Gini*

Decision Tree Algoritma *Index Gini* merupakan sebuah pohon keputusan yang terbentuk dari node pohon sehingga saling mempengaruhi faktor yang diuji dengan menggunakan nilai Gini Split terkecil sebagai *root node*.

Rumus *index gini* yang dinyatakan oleh Mandyartha, dkk (2015) terdapat pada persamaan 2.

$$Gini(S) = 1 - \sum_{k=1}^C p_k^2 \quad (2)$$

Keterangan:

K = Class atribut

C = Jumlah kelas pada variabel Y

P_k = Jumlah kelas dalam atribut K terhadap kelas pada variabel Y

3. HASIL DAN PEMBAHASAN

Penelitian menghasilkan analisis perbandingan menggunakan metode *Naive Bayes* dan algoritma *Index Gini*. Penelitian ini digunakan untuk mengetahui metode mana yang lebih utama dilakukan dalam memproses data dalam jumlah yang besar. Selain hal tersebut, penelitian ini bertujuan juga untuk mencari faktor apakah yang paling berpengaruh dalam penyalahgunaan narkoba. Hasil dari penelitian yang dilakukan sebagai berikut:

3.1. Implementasi *Naive Bayes* Menggunakan *Rapid Miner 5*

Dalam memprediksi waktu penggunaan pemakai narkoba di Panti Rehabilitasi Sinai Sukoharjo menggunakan metode *Naive Bayes* menggunakan *RapidMiner 5* dapat dilihat gambar 1.

SimpleDistribution

Distribution model for label attribute Penggunaan_Narkoba

Class Lama (0.340)

4 distributions

Class Tidak (0.660)

4 distributions

Gambar 1. Hasil *naive bayes* pada *Text View*

Pada gambar 1 menghasilkan model distribusi *naive bayes* nilai kelas LAMA (0.340) dan TIDAK (0.660). Gambar 2 merupakan data mentah yang digunakan untuk melakukan sebuah prediksi dengan membagi menjadi data training dan data testing.

Kepribadian	Lingkungan	Pendidikan	Jenis_Kelamin	Penggunaan_Narkoba
Coba-Coba	Masyarakat	Rendah	Laki-laki	Lama
Gangguan_Emosi	Teman	Rendah	Laki-laki	Tidak
Coba-Coba	Masyarakat	Rendah	Laki-laki	Tidak
Coba-Coba	Teman	Rendah	Laki-laki	Lama
Gangguan_Emosi	Teman	Rendah	Laki-laki	Tidak
Coba-Coba	Masyarakat	Rendah	Laki-laki	Lama
Coba-Coba	Masyarakat	Rendah	Laki-laki	Tidak
Gangguan_Emosi	Teman	Rendah	Laki-laki	Tidak
Gangguan_Emosi	Teman	Rendah	Laki-laki	Lama
Coba-Coba	Teman	Rendah	Laki-laki	Tidak
Coba-Coba	Teman	Rendah	Laki-laki	Tidak
Coba-Coba	Masyarakat	Rendah	Laki-laki	Tidak
Coba-Coba	Teman	Rendah	Laki-laki	Lama
Coba-Coba	Teman	Rendah	Laki-laki	Lama
Gangguan_Emosi	Teman	Rendah	Laki-laki	Lama
Coba-Coba	Masyarakat	Rendah	Laki-laki	Lama
Coba-Coba	Masyarakat	Rendah	Laki-laki	Tidak

Gambar 2. Data mentah

Hasil prediksi yang telah dieksekusi dengan jumlah data 300 sebagai data training dan 31 sebagai data testing terdapat pada gambar 3.

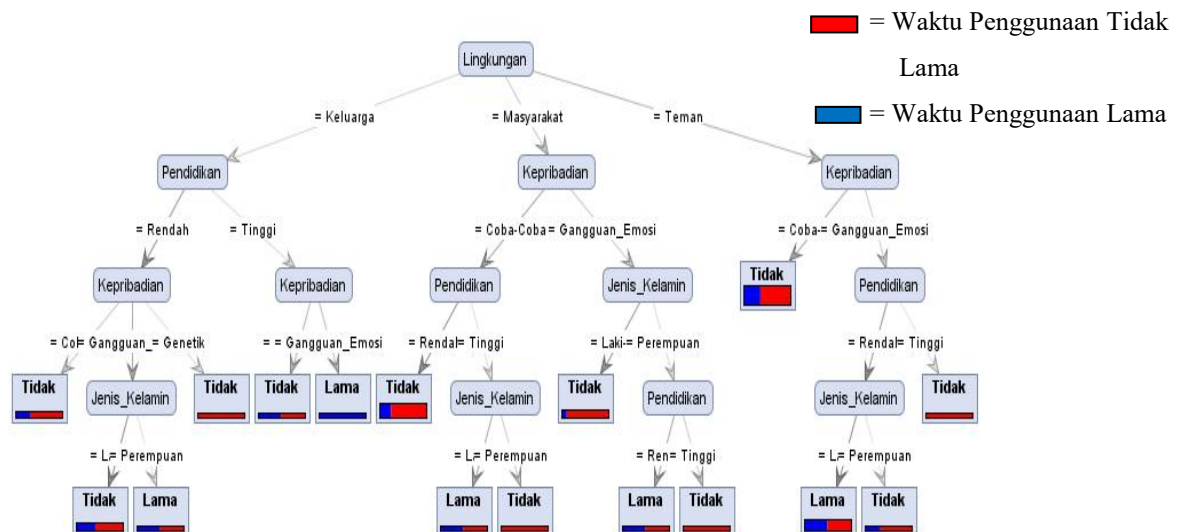
Row No.	confidence(Lama)	confidence(Tidak)	prediction(Penggunaan_Narkoba)	Kepribadian	Lingkungan	Pendidikan	Jenis_Kelamin
1	0.501	0.499	Lama	Gangguan_Emosi	Keluarga	Tinggi	Laki-laki
2	0.426	0.574	Tidak	Coba-Coba	Teman	Tinggi	Laki-laki
3	0.501	0.499	Lama	Gangguan_Emosi	Keluarga	Tinggi	Laki-laki
4	0.426	0.574	Tidak	Coba-Coba	Teman	Tinggi	Laki-laki
5	0.501	0.499	Lama	Gangguan_Emosi	Keluarga	Tinggi	Laki-laki
6	0.437	0.563	Tidak	Coba-Coba	Keluarga	Tinggi	Laki-laki
7	0.188	0.812	Tidak	Coba-Coba	Masyarakat	Rendah	Perempuan
8	0.347	0.653	Tidak	Coba-Coba	Teman	Rendah	Perempuan
9	0.408	0.592	Tidak	Gangguan_Emosi	Teman	Rendah	Perempuan
10	0.188	0.812	Tidak	Coba-Coba	Masyarakat	Rendah	Perempuan
11	0.388	0.612	Tidak	Coba-Coba	Keluarga	Rendah	Laki-laki
12	0.347	0.653	Tidak	Coba-Coba	Teman	Rendah	Perempuan
13	0.347	0.653	Tidak	Coba-Coba	Teman	Rendah	Perempuan
14	0.501	0.499	Lama	Gangguan_Emosi	Keluarga	Tinggi	Laki-laki
15	0.347	0.653	Tidak	Coba-Coba	Teman	Rendah	Perempuan
16	0.347	0.653	Tidak	Coba-Coba	Teman	Rendah	Perempuan
17	0.388	0.612	Tidak	Coba-Coba	Keluarga	Rendah	Laki-laki
18	0.408	0.592	Tidak	Gangguan_Emosi	Teman	Rendah	Perempuan

Gambar 3. Hasil prediksi penggunaan narkoba dengan metode *naive bayes*

3.2. Implementasi Algoritma *Index Gini* Menggunakan *Rapid Miner 5*

Memprediksi waktu penggunaan pemakai narkoba dengan metode *Decision Tree* Algoritma *Index Gini* menggunakan *RapidMiner 5* menghasilkan rancangan pada gambar 4 sebagai berikut.

Keterangan:



Gambar 4. Pohon keputusan metode *index gini*

Hasil gambar 4 menyatakan faktor lingkungan paling berpengaruh seseorang mengkonsumsi narkoba. Atribut Lingkungan menempati posisi *root node*. Internal

node pertama Lingkungan Keluarga yaitu atribut pendidikan. Internal node pertama Lingkungan Masyarakat dan Teman ditempati oleh atribut Kepribadian.

3.3. Implementasi Metode *Naive Bayes*

Salah satu data uji dijadikan contoh yaitu: kepribadian Gangguan Emosi, lingkungan Keluarga, pendidikan Tinggi, jenis kelamin Laki-laki. Apakah waktu penggunaan narkoba Lama ataukah Tidak?

Fakta dari contoh tersebut menunjukkan:

$$P(Y=Lama) = 112/331 = 0,338$$

$$P(Y=Tidak) = 219/331 = 0,662$$

Fakta:

$$P(X_1 = Gangguan Emosi | Y = Lama) = 44/112 = 0,393$$

$$P(X_1 = Gangguan Emosi | Y = Tidak) = 73/219 = 0,333$$

$$P(X_2 = Keluarga | Y = Lama) = 27/112 = 0,241$$

$$P(X_2 = Keluarga | Y = Tidak) = 55/219 = 0,251$$

$$P(X_3 = Tinggi | Y = Lama) = 8/112 = 0,071$$

$$P(X_3 = Tinggi | Y = Tidak) = 12/219 = 0,055$$

$$P(X_4 = Laki-laki | Y = Lama) = 106/112 = 0,946$$

$$P(X_4 = Laki-laki | Y = Tidak) = 197/219 = 0,900$$

HIMAP dihitung sebagai berikut:

$$P(X_1=Gangguan Emosi, X_2=Keluarga, X_3=Tinggi, X_4=Laki-laki | Y=Lama)$$

$$= \frac{44}{112} \times \frac{27}{112} \times \frac{8}{112} \times \frac{106}{112} \times 0,338 = 0,00215$$

$$P(X_1=Gangguan Emosi, X_2=Keluarga, X_3=Tinggi, X_4=Laki-laki | Y=Tidak)$$

$$= \frac{73}{219} \times \frac{55}{219} \times \frac{12}{219} \times \frac{197}{219} \times 0,662 = 0,00175$$

Nilai waktu penggunaan narkoba Lama $0,00215 > 0,00175$ Tidak, maka waktu penggunaan narkoba dengan fakta kepribadian Gangguan Emosi, lingkungan Keluarga, pendidikan Tinggi, jenis kelamin Laki-laki adalah Lama.

3.4. Implementasi dengan *decision tree* algoritma *index gini*

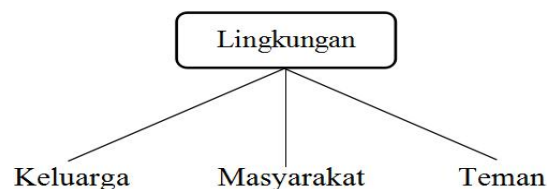
3.4.1. Perhitungan pada X1, X2, X3 X4 untuk membentuk *root node*.

Faktor yang mempunyai nilai *gini split* terkecil dijadikan *root node*.

Tabel 2. Hasil perhitungan *root node*

Faktor	Nilai <i>Gini Split</i>
X1=Kepribadian	0,446
X2=Lingkungan	0,434
X3=Pendidikan	0,446
X4=Jenis Kelamin	0,444

Hasil tabel 2 nilai *gini split* terkecil yaitu Lingkungan (0,434), sehingga terbentuk *root node* pada faktor Lingkungan.



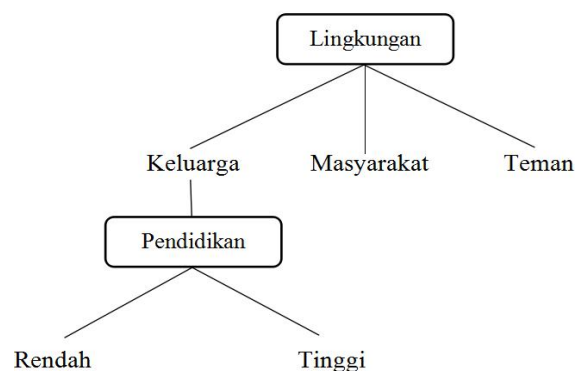
Gambar 5. *Root Node*

3.4.2. *Internal Node Pertama*.

Tabel 3. Hasil *internal node* pertama pada atribut Lingkungan Keluarga

Faktor	Nilai <i>Gini Split</i>
X1=Kepribadian	0,478
X3=Pendidikan	0,473
X4=Jenis Kelamin	0,486

Hasil tabel 3 nilai *gini split* terkecil yaitu Pendidikan (0,473), sehingga terbentuk *internal node* pada atribut Lingkungan Keluarga.



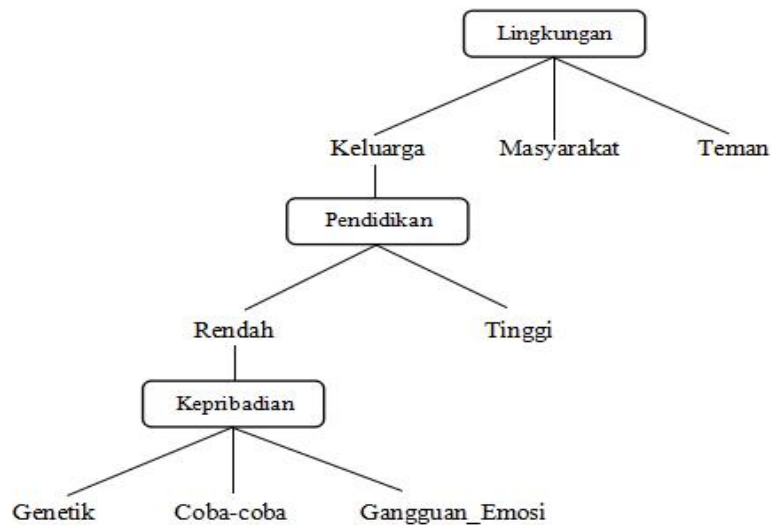
Gambar 6. *Internal Node* pertama atribut Pendidikan

3.4.3. Internal Node Kedua.

Tabel 4. Hasil pada faktor Lingkungan Keluarga, Pendidikan Rendah

Faktor	Nilai <i>Gini Split</i>
X1=Kepribadian	0,467
X4=Jenis Kelamin	0,475

Hasil tabel 4 nilai *gini split* terkecil yaitu Kepribadian (0,467), sehingga terbentuk *internal node* kedua di atribut Lingkungan Keluarga, Pendidikan Rendah.



Gambar 7. *Internal Node* kedua atribut Kepribadian

3.5. Hasil Perbandingan dari Metode

Hal ini digunakan untuk mendapatkan metode yang lebih akurat antara *Naive Bayes* dan *Decision Tree* Algoritma *Index Gini*. Hasil perbandingan sebagai berikut:

Tabel 5. Hasil Perbandingan

Nilai	<i>Naive Bayes</i>	<i>Index Gini</i>
<i>Accuracy</i>	65%	60%
<i>Precision</i>	65,66%	65,60%
<i>Recall</i>	98,48%	82,83%

Disimpulkan dari hasil tabel 5 *naive bayes* lebih akurat hasilnya dibandingkan dengan algoritma *Index Gini*.

3.6. Interpretasi Hasil Penelitian

Interpretasi hasil analisis dengan metode yang telah ditentukan dapat dijabarkan sebagai berikut:

3.6.1. Interpretasi Metode Naive Bayes

Metode tersebut memprediksi variabel yang mempengaruhi sebuah peluang. Hasil metode *naive bayes* terbukti waktu penggunaan narkoba “lama” sebesar 0,340 dan “tidak” sebesar 0,660. Waktu penggunaan narkoba ditentukan oleh atribut Jenis Kelamin “Laki-laki” ataupun “Prempuan” dan atribut Lingkungan “Keluarga”, “Masyarakat” maupun “Teman” menyatakan bahwa waktu penggunaan pemakai narkoba cenderung “Tidak”.

3.6.2. Interpretasi Algoritma *Index Gini*

Hasil algoritma membentuk sebuah pohon keputusan dimana atribut lingkungan merupakan faktor yang paling berpengaruh karena lingkungan sebagai *root node*. Hasil pohon keputusan pada atribut Lingkungan Keluarga maka *internal node* yang pertama yaitu atribut Pendidikan. Atribut lingkungan “Keluarga” dan Pendidikan “Tinggi” menempati *internal node* kedua adalah atribut Kepribadian.

4. PENUTUP

Hasil analisa dengan metode *decision tree* dalam penelitian ini ditarik kesimpulan bahwa atribut Lingkungan sangat berpengaruh dalam faktor penyalahgunaan narkoba sebagai *root node*. Sehingga dapat dilakukan penyuluhan mengenai dampak negatif dari pemakaian narkoba pada faktor Lingkungan.

Hasil prediksi menunjukkan bahwa *naive bayes* memiliki nilai *accuracy* 65%, *precision* 65,66%, *recall* 98,48%. Hasil *index gini* memiliki nilai *accuracy* 60%, *precision* 65,60%, *recall* 82,83%. Berdasarkan hasil tersebut metode *naive bayes* lebih akurat digunakan dibandingkan metode *decision tree* algoritma *index gini*. Dapat disimpulkan bahwa metode yang digunakan bisa memprediksi faktor yang paling berpengaruh dalam penelitian ini. Namun diharapkan untuk penelitian selanjutnya dapat menggunakan metode lain agar dapat dibandingkan hasilnya dan perlu adanya penambahan jumlah variabel yang sesuai dengan penelitian tersebut.

DAFTAR PUSTAKA

- Al-Bayaty, B.F.Z., & Joshi, S. (2016). Comparative Analysis between Naive Bayes Algorithm and Decision Tree to Solve WSD Using Empirical Approach, *Lecture Notes on Software Engineering*, 4(1).
- Iswanto, M.H., Permanasari, A.E., & Nugroho, H.A. (2015). Pemanfaatan Teknik Data Mining Untuk Diagnosis Penyakit *Tuberculosis* (TBC), Seminar Nasional Teknologi dan Multimedia, ISSN : 2302-3805.
- Mandyartha, E.P., Kurniawan, M., & Perdana, R.S. (2015). Identifikasi Sel Darah Merah Bertumpuk Menggunakan Pohon Keputusan *Fuzzy* Berbasis *Gini Index*, *Jurnal Buana Informatika*, 6(1), 51-62.
- Nasution, Marnis. (2014). Implementasi Data Mining Rough Set dalam Menentukan Tingkat Kerusakan Alat dan Bahan Kimia (Studi Kasus di Laboratorium Resep SMK S-16 Farmasi Bengkulu), *J.Informatika*, 2(1).
- Nugroho, Y.S., & Haryati, S.N. (2015). Klasifikasi dan Klastering Penjurusan Siswa SMA Negeri 3 Boyolali, *Khazanah Informatika*, 1(1), ISSN : 2477-698X.
- Ricardo, Paul. (2010). Upaya Penanggulangan Penyalahgunaan Narkoba oleh Kepolisian (Studi Kasus Satuan Narkoba Polres Metro Bekasi), *Jurnal Kriminologi Indonesia*, 6(3), 232-245.